

# Universality of generalized bunching and efficient assessment of Boson Sampling

V. S. Shchesnovich

Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, SP, 09210-170 Brazil

It is found that identical bosons (fermions) show generalized bunching (antibunching) property in linear networks: The absolute maximum (minimum) of probability that all  $N$  input particles are detected in a subset of  $\mathcal{K}$  output modes of any nontrivial linear  $M$ -mode network is attained *only* by completely indistinguishable bosons (fermions). For fermions  $\mathcal{K}$  is arbitrary, for bosons it is either (i) arbitrary for only classically correlated bosons or (ii) satisfies  $\mathcal{K} \geq N$  (or  $\mathcal{K} = 1$ ) for arbitrary input states of  $N$  particles. The generalized bunching allows to certify in a *polynomial* in  $N$  number of runs that a physical device realizing Boson Sampling with an *arbitrary* network operates in the regime of full quantum coherence compatible *only* with completely indistinguishable bosons. The protocol needs *only polynomial* classical computations for the standard BosonSampling, whereas an *analytic formula* is available for the Scattershot version.

PACS numbers: 42.50.St, 03.67.Ac, 42.50.Ar

*Introduction.*— Optical networks with photons have become a growing research field with potential application in quantum computing [1–3]. Boson Sampling (BS) idea [2], a non-universal but near-future feasible device aiming at the Extended Church-Turing thesis (ECT), followed by spectacular experiments with optical networks of growing size and number of photons [4–10], can be a way for benchmark demonstration of quantum supremacy. A BS device with a random but known  $M$ -mode network,  $N \sim 30$  single photons at known input modes for  $M \gg N^2$  would achieve this ultimate goal [2]. However, the very computational complexity of BS [11, 12] requires an exponential in  $N$  number of runs of such a device and computation of an exponential number (at least  $O(N^N)$ ) of classically hard permanents to prove BS by comparing an output distribution with theoretical probabilities. Quantum supremacy demonstration thus faces a big challenge of maintaining the  $N$ th order quantum coherence in a BS device with  $N \sim 30$  for an exponential number of runs. Though a distribution claimed to simulate BS, as the uniform one [13], can be exposed with only polynomial number of runs of a device [14] (see also Ref. [9]), finding such a protocol for a given sampler, e.g., the one with distinguishable particles, is a hard open problem.

The proof of BS being exponential both in the number of runs and computations does not prevent efficient verification of the very source of quantum supremacy of BS, i.e., the full  $N$ th-order quantum coherence in a device with  $N$  bosons, where unwanted distinguishability, photon losses, and higher photon numbers are the leading adversary factors in optical setups [4–10]. Such an assessment of BS may require only a polynomial number of runs and polynomial classical computations. Zero-transmission laws in the Fourier network [15] and statistical benchmarking [16] probed this path, but an assessment protocol applicable to an *arbitrary network* is an open problem. The test of Ref. [15] does verify the  $N$ th order coherence but only in a single Fourier network (not posing a threat to the ECT), whereas the statistical method of Ref. [16] can only assess the second-order

coherence. Already with relatively small networks experimentalists have to resort to either Bayesian methods or circumstantial evidence [8–10].

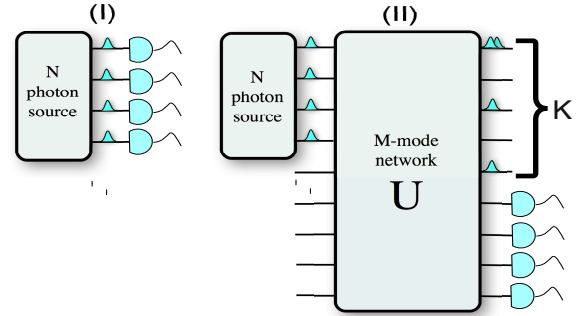


FIG. 1: (Color online) Assessment protocol for a BS device. Step (I): the source is checked for output with only one particle per mode. Step (II), the statistics of all  $N$  input particles to land in  $\mathcal{K}$  output modes of a  $M$ -mode network  $U$  is gathered by on-off type detectors in complementary  $L = M - \mathcal{K}$  modes: The maximum probability is attained *only* by completely indistinguishable bosons (and is close to 1 for  $NL \ll M$ ).

*Generalized bunching and efficient assessment of BS.*— We call a protocol that uses only polynomial classical computations to certify in a polynomial number of runs that a BS device operates at the full quantum  $N$ th-order coherence compatible *only* with completely indistinguishable bosons an efficient assessment protocol for BS. Efficiency *requires* an *independent* network certification: An efficient assessment of BS not using certification of a network matrix in a regime different from BS allows for loopholes, due to a combined effect of network and distinguishability errors (see the Appendix). This conclusion agrees with the results of Refs. [13, 14]. An assessment protocol for a BS device with an *arbitrary network* can only be based on a network-independent (universal) *absolute maximum or minimum* of some probability attained *only* by completely indistinguishable bosons [46]. The

discovered generalized boson bunching is such a property: The absolute maximum of probability of detecting *all*  $N$  input particles in  $\mathcal{K}$  output modes of an arbitrary nontrivial  $M$ -mode linear network is attained *only* by completely indistinguishable bosons (i) over arbitrary input states of particles for  $\mathcal{K} \geq N$  (and  $\mathcal{K} = 1$ ) or (ii) for arbitrary  $1 \leq \mathcal{K} \leq M - 1$  over only classically-correlated bosons. The minimum probability is attained by completely indistinguishable fermions, with that of distinguishable particles lying in the middle between the two. The generalized bunching/antibunching is found by exploiting the discovered equivalence between probability of detecting all input particles in a subset of output modes of a linear network and an eigenvalue problem for positive semi-definite (p.s.d.) Hermitian matrix (Eq. (7) below).

The generalized bunching can be viewed as an  $N$ -particle generalization of the Hong-Ou-Mandel (HOM) effect [17] to arbitrary  $M$ -mode networks. Known generalizations of the HOM effect to many-particle multi-mode setups are reported for special Bell multiport networks [18–20], whereas bunching to a single mode of a random network [21] has probability decaying as  $\sim e^{-N}$  for  $M \geq N^2$  [22]. A similar effect is boson tendency to form clouds in output modes in *some networks* [8]. *In contrast, the generalized bunching can be observed in a polynomial in  $N$  number of experimental runs in an arbitrary (nontrivial) quantum network.* Though in the dilute limit,  $M \geq N^2$ , the generalized bunching/antibunching effect wanes as  $N \rightarrow \infty$ , it nevertheless allows for an efficient assessment protocol of a BS device, sketched in Fig. 1, requiring, besides only a polynomial number of runs of a device, polynomial classical computations (see below). Moreover, analytical results are derived for the Scattershot version of BS [23].

*Description of partially distinguishable identical particles in a linear network.*— We consider an  $M$ -mode linear quantum network with  $N$  single identical particles in an arbitrary internal state at input modes  $k_1, \dots, k_N$  (ordered products of operators are assumed in case of fermions)

$$\rho = \sum_i q_i |\Psi_i\rangle \langle \Psi_i|, \quad |\Psi_i\rangle = \sum_{\mathbf{j}} C_{\mathbf{j}}^{(i)} \prod_{\alpha=1}^N a_{k_{\alpha}, j_{\alpha}}^{\dagger} |0\rangle, \quad (1)$$

where  $q_i \geq 0$ ,  $\sum_i q_i = 1$ ,  $\mathbf{j} = (j_1, \dots, j_N)$ ,  $\sum_{\mathbf{j}} |C_{\mathbf{j}}^{(i)}|^2 = 1$ , and a mode operator  $a_{k,j}^{\dagger}$  (and below  $b_{k,j}^{\dagger}$ ) creates a particle in an input (output) mode  $k$  and an internal basis state  $|j\rangle \in \mathcal{H}$  (e.g., a basis function of spectral shape of a photon). A unitary network with matrix  $U$ , Fig. 1(b), relates the input  $a_{k,j}$  and output  $b_{l,j}$  modes:  $a_{k,j}^{\dagger} = \sum_{l=1}^M U_{k,l} b_{l,j}^{\dagger}$ . The internal state of identical particles, defined as

$$\rho^{(int)} = \sum_i q_i |\psi_i\rangle \langle \psi_i|, \quad |\psi_i\rangle \equiv \sum_{\mathbf{j}} C_{\mathbf{j}}^{(i)} \prod_{\alpha=1}^N |j_{\alpha}\rangle, \quad (2)$$

governs their behavior in a linear network. Symmetry properties of  $\rho^{(int)}$  under the symmetric group  $\mathcal{S}_N$  play

the key role [24–29], e.g., particles of one species with  $\rho^{(int)}$  antisymmetric under permutations emulate behavior of the other species [30, 31]. The probability formula of an output configuration  $\mathbf{m} = (m_1, \dots, m_M)$  [24, 25] (applicable also to fermions, see the Appendix) reads

$$\hat{p}(\mathbf{m}) = \frac{1}{\prod_{l=1}^M m_l!} \sum_{\tau, \sigma \in \mathcal{S}_N} J(\tau\sigma^{-1}) \prod_{\alpha=1}^N U_{k_{\tau(\alpha)}, l_{\alpha}}^* U_{k_{\sigma(\alpha)}, l_{\alpha}}, \quad (3)$$

where  $l_1, \dots, l_N$  are output modes with multiplicities  $(m_1, \dots, m_M)$ , and a complex-valued function  $J(\sigma)$  of  $\sigma \in \mathcal{S}_N$  is defined as

$$J(\sigma) = \varepsilon(\sigma) \text{Tr}(\rho^{(int)} P_{\sigma}), \quad \varepsilon(\sigma) = \begin{cases} 1, & \text{Bosons,} \\ \text{sgn}(\sigma), & \text{Fermions,} \end{cases} \quad (4)$$

where  $P_{\sigma} \prod_{\alpha=1}^N |j_{\alpha}\rangle = \prod_{\alpha=1}^N |j_{\sigma^{-1}(\alpha)}\rangle$  is an operator representation of  $\sigma$  in the Hilbert space  $\mathcal{H}^{\otimes N}$ .

Identical particles are called completely indistinguishable if  $\rho^{(int)}$  is symmetric under permutations,  $J^{(id)}(\sigma) = \varepsilon(\sigma)$  (e.g., particles in the same internal state), whereas particles with orthogonal internal states are distinguishable,  $J^{(d)}(\sigma) = \delta_{\sigma, I}$  (see also Refs. [25, 31]). The trace of  $\rho^{(int)}$  in the symmetric subspace  $\mathcal{S}_N \mathcal{H}^{\otimes N}$  with  $\mathcal{S}_N = (1/N!) \sum_{\sigma} P_{\sigma}$ , i.e.,  $d(J) \equiv \text{Tr}\{\mathcal{S}_N \rho^{(int)}\} = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \varepsilon(\sigma) J(\sigma)$ , is a suitable measure of the partial indistinguishability [28, 32] and also in other quantum information applications of identical particles [33].

Note that  $J(\sigma)$  of Eq. (4) is a p.s.d. function of  $\sigma \in \mathcal{S}_N$ , i.e., for any complex-valued function  $z(\sigma)$  we have  $\sum_{\sigma_1, \sigma_2} z^*(\sigma_1) J(\sigma_1 \sigma_2^{-1}) z(\sigma_2) \geq 0$ , while  $J(I) = 1$  for the identity permutation  $I$ . Therefore, there is a factorizing function  $\theta(\sigma)$  such that (see the Appendix)

$$J(\sigma) = \sum_{\tau \in \mathcal{S}_N} \theta^*(\tau) \theta(\tau\sigma), \quad \sum_{\sigma \in \mathcal{S}_N} |\theta(\sigma)|^2 = 1. \quad (5)$$

Importantly, *any* normalized p.s.d. function  $J(\sigma)$  corresponds to an input state of single particles (see the Appendix), i.e., the whole set of input states with  $N$  single particles can be equivalently represented by the whole convex set of normalized p.s.d. functions  $\mathcal{S}_N \rightarrow \mathbb{C}$ .

*Probability to detect all particles in  $\mathcal{K}$  output modes.*— The probability for all  $N$  input particles to gather in  $\mathcal{K}$  output modes, say  $1, \dots, \mathcal{K}$  as in Fig. 1(b), reads  $p_N(J) = \sum_{\mathbf{m}} \hat{p}(\mathbf{m})$  where  $\mathbf{m} = (m_1, \dots, m_{\mathcal{K}}, 0, \dots, 0)$ . Defining an  $N$ -dimensional p.s.d. Hermitian matrix  $H$ , built from the submatrix of  $U$  on the rows  $k_1, \dots, k_N$  and columns  $1, \dots, \mathcal{K}$ , and the corresponding  $(N!)$ -dimensional Schur power matrix  $\Pi(H)$  indexed by elements of  $\mathcal{S}_N$ ,

$$H_{\alpha, \beta} \equiv \sum_{l=1}^{\mathcal{K}} U_{k_{\alpha}, l} U_{k_{\beta}, l}^*, \quad \Pi_{\sigma, \tau}(H) = \prod_{\alpha=1}^N H_{\sigma(\alpha), \tau(\alpha)}, \quad (6)$$

we obtain  $p_N(J)$  in the form (see the Appendix)

$$p_N(J) = \sum_{\sigma \in \mathcal{S}_N} J(\sigma) \Pi_{I,\sigma} = \sum_{\tau, \tau' \in \mathcal{S}_N} \theta^*(\tau) \Pi_{\tau, \tau'} \theta(\tau'). \quad (7)$$

Therefore,  $p_N(J)$  is a convex combination of eigenvalues of  $\Pi(H)$ . For the completely indistinguishable particles  $\theta^{(id)}(\sigma) = \varepsilon(\sigma)/\sqrt{N!}$ , whereas for distinguishable particles  $\theta^{(d)}(\sigma) = \delta_{\sigma, I}$ . Eq. (7) gives for these ideal cases:

$$p_N^{(B)} = \text{per}(H), \quad p_N^{(F)} = \det(H), \quad p_N^{(d)} = \prod_{\alpha=1}^N H_{\alpha, \alpha}, \quad (8)$$

for bosons, fermions, and classical particles, respectively. Well-known inequalities (see the Appendix) result in the following order

$$p_N^{(F)} \leq p_N^{(d)} \leq p_N^{(B)}. \quad (9)$$

Eq. (9) seem to suggest that  $p_N^{(B)}$  and  $p_N^{(F)}$  are the absolute maximum and minimum of  $p_N(J)$ . To establish such a property one has to show that they are the *unique maximum (minimum) eigenvalues* of  $\Pi(H)$ . A famous result of Schur [35] states that the smallest eigenvalue of  $\Pi(H)$  is  $\det(H)$ , hence, the generalized antibunching is an universal attribute of the completely indistinguishable fermions (a unique minimum for  $\mathcal{K} \geq N$ ). On the other hand, the maximum eigenvalue of  $\Pi(H)$  is generally unknown. The permanent-on-top conjecture (POT), stating that universally it is  $\text{per}(H)$ , proven for  $N \leq 3$  [36], has turned out false for  $N \geq 5$  [37].

Thus, conditions on input state and/or network are needed to ensure the maximum probability being attained only by the completely indistinguishable bosons. For  $\mathcal{K} = 1$  Eq. (7) gives  $p_N(J) = d(J)N! \prod_{\alpha=1}^N |U_{k_\alpha, l}|^2$  with the maximum for an arbitrary network at  $d(J) = 1$  or  $J(\sigma) = 1$  (see also Refs. [21, 28, 32]). Numerical simulations [47] with random p.s.d. Hermitian matrices reveal that  $p_N^{(B)} = \text{per}(H)$  is not the maximum probability *only* for  $N \geq 5$  particles in  $2 \leq \mathcal{K} \leq N - 1$  output modes (such a state of  $N \geq 5$  bosons is necessarily a state of non-classically correlated particles, see below). More importantly, when  $\mathcal{K} \geq N$ , a *unique eigenvector*  $\theta_B^{(id)}(\sigma) = 1/\sqrt{N!}$  corresponds to  $\text{per}(H)$ , i.e., the maximum probability of  $N$  particles to be detected in  $\mathcal{K} \geq N$  output modes is attained *only* by the completely indistinguishable bosons. Thus, bosons show the generalized bunching property in  $\mathcal{K} \geq N$  output modes.

For only classically correlated bosons, a unique maximum of  $p_N(J)$  is attained only by the completely indistinguishable ones for all  $1 \leq \mathcal{K} \leq M - 1$ . Indeed, a classically-correlated internal state can be expressed as a convex combination of pure states, i.e.,  $\rho^{(int)} = \sum_j \nu_j |\phi_{j_1}\rangle\langle\phi_{j_1}| \otimes \dots \otimes |\phi_{j_N}\rangle\langle\phi_{j_N}|$  for some arbitrary states  $|\phi_{j_\alpha}\rangle \in \mathcal{H}$  and  $\nu_j > 0$ ,  $\sum_j \nu_j = 1$ . The corresponding  $J$ -function reads

$$J^{(cc)}(\sigma) = \sum_j \nu_j \prod_{\alpha=1}^N \langle\phi_{j_{\sigma(\alpha)}}|\phi_{j_\alpha}\rangle. \quad (10)$$

Setting  $G_{\alpha, \beta}^{(j)} \equiv \langle\phi_{j_\beta}|\phi_{j_\alpha}\rangle$ , we get from Eqs. (6), (7), and (10)

$$p_N(J^{(cc)}) = \sum_j \nu_j \text{per}(H \cdot G^{(j)}), \quad (11)$$

where the dot stands for the Hadamard (by-element) product. Thus, the permanent version of Oppenheim's inequality [38], stating that for two p.s.d. Hermitian matrices  $H$  and  $G$  (for  $G_{\alpha, \alpha} = 1$ ),  $\text{per}(H \cdot G) \leq \text{per}(H)$  would imply the claimed result in this case. Using matrices  $H$  that violate the POT conjecture, it was checked that  $\text{per}(H \cdot G) < \text{per}(H)$  for any  $N$  random states  $|\phi_1\rangle, \dots, |\phi_N\rangle$ , with at least two linearly independent.

*Assessment protocol of a BS device.*— The average  $\langle p_N(J) \rangle$  over Haar-random networks gives an idea of quantitative features of the generalized bunching/antibunching effect. By the unitary invariance of the Haar measure, the average probability  $\langle p_N^{(B, F)} \rangle$  depends only on the ratio of considered output configurations (see the Appendix):

$$\langle p_N^{(B, F)} \rangle = \frac{\mathcal{K}(\mathcal{K} \pm 1) \cdot \dots \cdot (\mathcal{K} \pm N \mp 1)}{M(M \pm 1) \cdot \dots \cdot (M \pm N \mp 1)}, \quad (12)$$

here (and below) the upper (lower) signs stand for bosons (fermions). For distinguishable particles there is no exact result, but for  $M \gg 1$  it can be shown that (see the Appendix)

$$\langle p_N^{(d)} \rangle = \left( \frac{\mathcal{K}}{M} \right)^N \left[ 1 + O\left( \frac{N^2}{\mathcal{K}M} \right) \right]. \quad (13)$$

The average probability ratio becomes

$$\frac{\langle p_N^{(B, F)} \rangle}{\langle p_N^{(d)} \rangle} = \left[ 1 + O\left( \frac{N^2}{\mathcal{K}M} \right) \right] \prod_{i=1}^{N-1} \frac{1 \pm i/\mathcal{K}}{1 \pm i/M}. \quad (14)$$

For  $NL \ll M$ , where  $L = M - \mathcal{K}$ , the detection probability is close to 1:  $\langle p_N^{(B, F)} \rangle = \prod_{l=0}^{N-1} [1 - L/(M \pm l)] = 1 - O(LN/M)$ , whereas the r.h.s. of Eq. (14) gives  $\langle p_N^{(B, F)} \rangle / \langle p_N^{(d)} \rangle \approx 1 \pm LN(N-1)/(2M^2)$ . In this case one needs  $R \gg M^4/(N^4 L^2)$  runs for the ratio (14) to surpass the statistical error  $O(1/\sqrt{R})$  in experimental data. But the ratio in Eq. (14) is attained *only* by the completely indistinguishable bosons (fermions), hence, it is a *reliable witness*, detectable in polynomial number of runs, of their complete indistinguishability during propagation, i.e., that no decoherence process has contributed to output statistics. Therefore, we have an efficient protocol for assessment of a BS device with an *arbitrary* network, Fig. 1. The only known protocol for a BS device with an arbitrary network [14], experimentally verified [9], discriminates the BS and uniform distributions. Our protocol discriminates against *all* other than BS samplers, which are physically realizable with particles in a linear network, including the classical ( $\mathcal{M}_A$ ), the fermion ( $\mathcal{F}_A$ ),

the random-classical ( $\mathcal{B}_A$ ) samplers of Ref. [14], and the random-phase bosons [15, 16].

The assessment protocol has two stages. At stage (I), Fig. 1(I), by using photon-number resolving detectors (e.g., by cascading bucket detectors [39]) one checks that sources produce  $N$  single photons. At stage (II), Fig. 1(II), employing the universality of generalized bunching, one verifies experimental statistics against the probability  $p_N^{(B)}$  (8) using  $L = M - \mathcal{K}$  bucket detectors.

The protocol requires *only one* matrix permanent  $p_N^{(B)} = \text{per}(H)$  of  $H$  in Eq. (6) (a single set of  $\mathcal{K}$  modes is used) to an error  $\epsilon = O(N^{-\kappa})$  for some  $\kappa > 0$  (statistical error in experimental data for a polynomial number of runs). For  $N \gg 1$  in the dilute limit  $M = O(N^{2+\delta})$  with  $\delta > 0$  and  $L = O(N)$  it can be shown (using that we select  $\mathcal{K}$  modes arbitrarily) that *only polynomial in  $N$  computations  $\mathcal{C}_N$*  are required (see the Appendix; in this case  $\langle p_N^{(B)} \rangle = 1 - O(N^{-\delta})$  and  $\langle p_N^{(B)} - p_N^{(d)} \rangle = O(N^{-1-2\delta})$ ). One can estimate that, on average over all choices of  $\mathcal{K}$  output modes,  $\mathcal{C}_N = O(N^{\frac{\kappa+1}{\delta}})$  (e.g., setting  $\kappa = 2[1 + \delta]$  allows one to distinguish the quantum and classical cases).

The protocol applies also to the Scattershot BS [23], recently experimentally tested [10], which uses heralded single photons in  $N$  random input modes in each run. The probability describing experimental statistics of a Scattershot BS is well approximated by  $\langle p_N^{(B)} \rangle$  (12) already for a few hundred runs of such a device (see Fig. 18 below), i.e., *no computations required*.

Stage (I) is designed to expose all attempts to bypass the universality property using inputs with variable number of particles per mode (not required under certified input). Indeed, an input with any distribution of the completely indistinguishable bosons between  $M$  input modes

$$\rho = \sum_{\mathbf{n}} p_{\mathbf{n}} |\mathbf{n}\rangle \langle \mathbf{n}|, \quad p_{\mathbf{n}} \geq 0, \quad \sum_{\mathbf{n}} p_{\mathbf{n}} = 1, \quad (15)$$

where  $n_1 + \dots + n_M = N$ , has the Haar-average probability equal to  $\langle p_N^{(B)} \rangle$  (see the Appendix). For example, take the random-phase bosons of Refs. [15, 16], i.e., an input where each boson is in a coherent superposition of  $S$  input modes (and in an internal state  $|\phi\rangle$ ) described by operator  $A(\theta) = S^{-\frac{1}{2}} \sum_{j=1}^S e^{i\theta_j} a_{k_j, \phi}$  with random phases  $\theta_1, \dots, \theta_S$  [15, 16]. The density matrix of this input reads

$$\begin{aligned} \rho_s &= \frac{(S-1)! S^N}{(S+N-1)!} \prod_{j=1}^S \int_0^{2\pi} \frac{d\theta_j}{2\pi} [A^\dagger(\theta)]^N |0\rangle \langle 0| [A(\theta)]^N \\ &= \frac{N!(S-1)!}{(S+N-1)!} \sum_{\mathbf{n}} |\mathbf{n}\rangle \langle \mathbf{n}|, \end{aligned} \quad (16)$$

where  $\mathbf{n} = (n_{k_1}, \dots, n_{k_S})$ ,  $n_{k_1} + \dots + n_{k_S} = N$ . A source of  $\rho_s$  with  $S = N$  is exposed at stage (I) by a vanishing probability of an input with one particle per occupied mode, for  $N \gg 1$  scaling as  $\sim 4^{-N}$ . Stage (I) exposes

also the sampler  $\mathcal{B}_A$  of Ref. [14], a “mockup distribution of BS” physically realized by distributing  $N$  uncorrelated distinguishable particles randomly over  $N$  input modes, with the probability of such a particle to land in an output mode  $l$  being  $p(l) = \frac{1}{N} \sum_{\alpha=1}^N |U_{k_\alpha, l}|^2$ . The probability of single particles at input is  $N!/N^N \sim e^{-N}$ .

$N$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$L$	2	2	3	4	5	5	6	7	7	8	9	9	10	11	11	12	13	14
$M$	5	8	13	18	25	32	41	50	61	72	85	98	113	128	145	162	181	200

TABLE I: Network size  $M$  and  $L = M - \mathcal{K}$  as functions of  $N$ . Here  $\mathcal{K}$  is selected by maximizing the ratio of Eq. (14) under the condition that  $\langle p_N^{(B)} \rangle \geq 0.25$  (note that  $\mathcal{K} \geq N$ ).

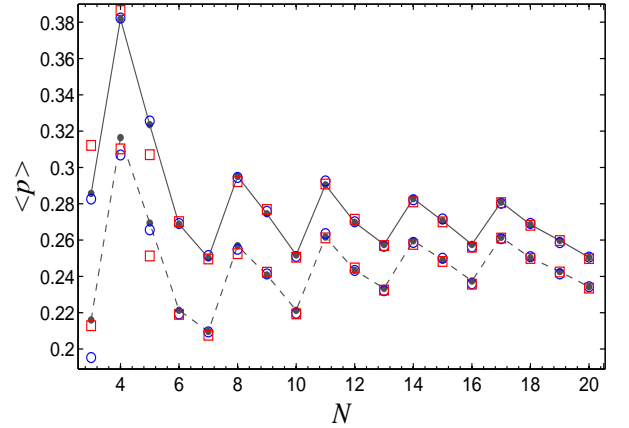


FIG. 2: (Color online) The analytical quantum average probability  $\langle p_N^{(B)} \rangle$  Eq. (12) (dots on the solid line) and the approximation Eq. (13) of the classical one  $\langle p_N^{(d)} \rangle$  (dots on the dashed line) vs. numerical averaging with 1000 Haar-random networks (circles). The squares give the quantum and classical probability for the Scattershot BS (estimated with 500 runs) with a randomly chosen network for each value of  $N$ .

Fig. 18 gives numerical results, where  $M = \lfloor \frac{1}{2} N^2 \rfloor$  (integer part), and  $L = M - \mathcal{K}$  is obtained by maximizing the ratio (14) for  $\langle p_N^{(B)} \rangle \geq 0.25$  (see Table I).

The distinguishability error  $1 - d(J) \approx (1 - F)(N - 1)$  [32] in a BS device ( $F$  is the mean fidelity of indistinguishability of photons) can be assessed from an experiment using the first-order approximations (see the Appendix):

$$\begin{aligned} p_N^{(B)} - p_N(J) &\approx \frac{1 - d(J)}{N - 1} \left[ N p_N^{(B)} - \frac{d}{dx} \text{per}\{H(x)\}_{x=1} \right], \\ \langle p_N^{(B)} \rangle - \langle p_N(J) \rangle &\approx [1 - d(J)] \frac{N}{M} \langle p_{N-1}^{(B)} \rangle, \end{aligned} \quad (17)$$

where  $H_{\alpha, \beta}(x) \equiv \delta_{\alpha, \beta} H_{\alpha, \beta} x + (1 - \delta_{\alpha, \beta}) H_{\alpha, \beta}$  ( $\text{per}\{H(x)\}$  is a polynomial in  $x$  of order  $N$ ). The second law is valid for  $M \gg N^2$  and applies to the Scattershot BS.

Difference between an experimental and the theoretical probability  $p_N^{(B)} - p_N^{(exp)}$  would reflect presence of other

errors in a device. How the BS regime is affected by errors in a network matrix can be estimated beforehand. The aim of such a certification, besides eliminating the possibility of loopholes (see the Appendix), is to guarantee that network errors would be below an acceptable level in the BS regime. The theoretical basis is provided in Refs. [40, 41]. Besides an unwanted distinguishability and network matrix errors, errors reported in BS experiments [4–10] include higher-order photon numbers, estimated at stage (I) of our protocol, and the photon losses, which can be directly estimated at a network output (see also Ref. [42]). A BS device with a fixed-ratio of lost photons is believed to be hard to simulate on a classical computer [2, 43] with some progress in proof [44]. Since a lossy linear  $M$ -mode network is equivalent to an  $2M$ -mode unitary one, with one half of output modes being inaccessible “loss channels” (see the Appendix), our assessment protocol applies also to linear lossy networks. In this case in Eq. (6) the proper non-unitary network matrix  $U$  must be used. It can be experimentally characterized with only classical coherence [45].

*Conclusion.*— We have discovered the generalized boson bunching and fermion antibunching in linear networks and proposed an assessment protocol for BS verifying in a *polynomial* number of experimental runs that a BS device with a random linear network operates at the full  $N$ th-order quantum coherence compatible only with  $N$  completely indistinguishable bosons, i.e., the very physical origin of its quantum supremacy. The protocol requires *only polynomial* classical computations for the standard version of BS, whereas for the Scattershot version (with better prospects for scalability) *analytical* results are available. In general terms, the generalized bunching is a generalization of the famous HOM effect, revealing the complete indistinguishability of  $N$  photons in an arbitrary (nontrivial)  $M$ -mode network, which may find other applications whenever linear quantum networks are used.

*Acknowledgement.*— This work was supported by the CNPq (Brazil) under grants 304986/2011-9 and 304129/2015-1.

- 
- [1] E. Knill, R. Laflamme, and G. J. Milburn, *Nature* **409** (2001) 46.
  - [2] S. Aaronson and A. Arkhipov, *Theory of Computing* **9**, 143 (2013).
  - [3] J. Huh, G. G. Guerreschi, B. Peropadre, J. R. McClean, and A. Aspuru-Guzik, *Nat. Photonics* **9**, 615 (2015).
  - [4] M. A. Broome *et al*, *Science* **339**, 794 (2013);
  - [5] J. B. Spring *et al*, *Science*, **339**, 798 (2013).
  - [6] M. Tillmann *et al*, *Nature Photonics*, **7**, 540 (2013).
  - [7] A. Crespi *et al*, *Nature Photonics*, **7**, 545 (2013).
  - [8] J. Carolan *et al*, *Nat. Photon.* **8**, 621 (2014).
  - [9] N. Spagnolo *et al*, *Nature Photonics*, **8**, 615 (2014).
  - [10] M. Bentivegna *et al*, *Sci. Adv.* **1**, e1400255 (2015).
  - [11] L. G. Valiant, *Theoretical Comput. Sci.*, **8**, 189 (1979).
  - [12] S. Aaronson, *Proc. Roy. Soc. London A*, **467**, 3393 (2011).
  - [13] C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert, *arXiv:1306.3995* [quant-ph].
  - [14] S. Aaronson and A. Arkhipov, *arXiv:1309.7460* [quant-ph].
  - [15] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Molmer, *Phys. Rev. Lett.* **113**, 020502 (2014).
  - [16] M. Walschaers *et al*, *arXiv:1410.8547* [quant-ph].
  - [17] C. K. Hong, Z. Y. Ou, and L. Mandel, *Phys. Rev. Lett.* **59**, 2044 (1987).
  - [18] Y. L. Lim and A. Beige, *New J. Phys.* **7**, 155 (2005).
  - [19] Z. Y. Ou, *Phys. Rev. A* **77**, 043829 (2008).
  - [20] M. C. Tichy, M. Tiersch, F. Mintert, and A. Buchleitner, *New Journal of Phys.* **14**, 093015 (2012).
  - [21] N. Spagnolo *et al*, *Phys. Rev. Lett.* **111**, 130503 (2013).
  - [22] A. Arkhipov and G. Kuperberg, *Geom. Topol. Monogr.*, **18**, 1 (2012).
  - [23] A. P. Lund *et al*, *Phys. Rev. Lett.* **113**, 100502 (2014).
  - [24] V. S. Shchesnovich, *Phys. Rev. A* **89**, 022333 (2014).
  - [25] V. S. Shchesnovich, *Phys. Rev. A* **91**, 013844 (2015).
  - [26] Z. Y. Ou, *Phys. Rev. A* **74**, 063808 (2006).
  - [27] H. de Guise, S.-H. Tan, I. P. Poulin, and B. C. Sanders, *Phys. Rev. A* **89**, 063819 (2014).
  - [28] M. C. Tichy, *Phys. Rev. A* **91**, 022316 (2015).
  - [29] P. P. Rohde, *Phys. Rev. A* **91**, 012307 (2015).
  - [30] J. C. F. Matthews *et al*, *Sci. Rep.* **3**, 1539 (2013).
  - [31] V. S. Shchesnovich, *Int. J. Quantum Inform.* **13**, 1550013 (2015).
  - [32] V. S. Shchesnovich, *Phys. Rev. A* **91**, 063842 (2015).
  - [33] M. Miková *et al*, *Phys. Rev. A* **87**, 042327 (2013).
  - [34] See the Supplemental material.
  - [35] I. Schur, *Math. Z.* **1**, 184 (1918).
  - [36] R. B. Bapat and V. S. Sunder, *Linear Algebra Appl.* **76**, 153 (1986).
  - [37] V. S. Shchesnovich, *Linear Algebra Appl.* **490**, 196 (2016).
  - [38] R. B. Bapat and V. S. Sunder, *Linear Algebra Appl.* **72**, 107 (1985).
  - [39] P. P. Rohde, *J. Opt. B: Quant. Semiclass. Opt.* **7** 82 (2005).
  - [40] A. Leverrier and R. García-Patrón, *Quant. Inform. Comput.* **15**, 489 (2015).
  - [41] A. Arkhipov, *Phys. Rev. A* **92**, 062326 (2015).
  - [42] V. S. Shchesnovich, *arXiv:1403.4459* [quant-ph].
  - [43] P. P. Rohde and T. C. Ralph, *Phys. Rev. A* **85**, 022332 (2012).
  - [44] S. Aaronson and D. J. Brod, *arXiv:1510.05245* [quant-ph].
  - [45] S. Rahimi-Keshari *et al*, *Optics Express* **21**, 13450 (2013).
  - [46] E.g., average  $N$ -fold detection probability of indistinguishable bosons is less than that of distinguishable particles [8], whereas a BS assessment protocol needs an absolute minimum over *all* inputs with partially distinguishable particles *in any non-trivial network*.
  - [47] Numerical simulations are limited to  $N \leq 7$  due to a large dimension of the Schur power matrix.

## Appendix

### I An efficient assessment protocol for BS on an uncertified network matrix allows for loopholes

Verification of Boson Sampling (BS) with a random network without knowledge of a network matrix is an ill-defined problem [1, 2]. In view of this, a natural question arises: Is it somehow possible to verify the full  $N$ th order quantum coherence compatible only with  $N$  completely indistinguishable bosons and a (special) network matrix simultaneously *in a polynomial number of runs* in the same BS regime on such a device? This may seem indeed possible due to presence of a high symmetry in a network matrix, leading to very distinct features in output distribution. Let us call such an assessment of BS *the holistic assessment*.

However, careful analysis reveals loopholes in an efficient assessment protocol if the latter is used as a holistic one. Let us analyze in detail an assessment protocol for BS based on the Fourier network [3] (a highly symmetric Bell multiport), where a large fraction of the output configurations are forbidden by symmetry for completely indistinguishable bosons [3, 4]. Let us recall the necessary details of the assessment protocol proposed in Ref. [3]. One considers Fourier network with  $M = N^p$  modes (where  $p \geq 2$ ), i.e., given by the matrix

$$F_{k,l}^{(M)} = \frac{1}{\sqrt{M}} \exp\left(i \frac{2\pi}{M} kl\right), \quad k, l = 1, \dots, M. \quad (18)$$

For single particles in the input modes  $1 \leq k_1, \dots, k_N \leq M$  with the following cyclic symmetry

$$k_{\alpha+1} = k_\alpha + N^{p-1} \quad (19)$$

where, obviously,  $1 \leq k_1 \leq N^{p-1}$ , only those output configurations  $1 \leq l_1, \dots, l_N \leq M$  which satisfy

$$\sum_{\alpha=1}^N l_\alpha = qN, \quad q \in \mathcal{N}, \quad (20)$$

i.e., have the sum of mode indices divisible by  $N$ , are realized with completely indistinguishable bosons in such a network [3]. Moreover, the number of realized output configurations is only a fraction  $1/N$  of the total number of all *a priori* possible ones, which allows efficient verification of the full  $N$ th order indistinguishability by counting the number of violations of the forbidden events, on the order of  $1 - 1/N$  of the total number of events [3].

The above described protocol is subject to loopholes, if it is understood as a holistic one, i.e. when no certification of a network matrix *in a different regime* is performed [23]. Loopholes appear due to a combined effect of network and distinguishability errors. Let us illustrate this by exposing a loophole in the following case:

$N = 2N_1$  and  $p = 2$  (for simplicity). Denote  $M_1 = N_1^2$ . Since no information on a network is assumed, instead of  $F^{(M)}$  (18) one cannot rule out the following one ( $P$  is a permutation)

$$U = P \left[ F^{(M_1)} \oplus F^{(M_1)} \right] P^\dagger \oplus P \left[ F^{(M_1)} \oplus F^{(M_1)} \right] P^\dagger, \quad (21)$$

where  $U$  is a block-structured matrix with two diagonal blocks, each containing two  $M_1$ -dimensional Fourier matrices (18) whose rows and columns are permuted by  $P$  in such a way that each odd (even) row and column in the two blocks of  $U$  of size  $2M_1$  corresponds to one and the same submatrix  $F^{(M_1)}$ .

Consider now  $N$  bosons divided into two groups of  $N_1$  completely indistinguishable bosons, whereas bosons from different groups being distinguishable (e.g., bosons in group  $i = 1, 2$  are in an internal state  $|\phi_i\rangle$  and  $\langle\phi_1|\phi_2\rangle = 0$ ). Assume that such bosons are launched into  $U$  (21) observing the cyclic symmetry (19), where the bosons in the first  $N_1$  input modes of  $U$  are from the same group (i.e., completely indistinguishable). Since  $N$  is even, the parity of an input mode  $k_\alpha$  of  $U$  is the same as that of  $k_1$ . Thus, the first (second) group of  $N_1$  completely indistinguishable bosons are launched in one of the two submatrices  $F^{(M_1)}$  of the respective block, in a similar cyclic symmetry as in Eq. (19) for its proper mode indices and  $N$  substituted by  $N_1$ . By Refs. [3, 4] and the block-structure of  $U$ , the allowed output configurations must satisfy conditions similar to Eq. (20) now for the proper output indices, say,  $1 \leq l_1^{(i)}, \dots, l_{N_1}^{(i)} \leq M_1$ , of that particular submatrix  $F^{(M_1)}$ , i.e.,

$$\sum_{\alpha=1}^{N_1} l_\alpha^{(i)} = q^{(i)} N_1, \quad q^{(i)} \in \mathcal{N}, \quad i = 1, 2. \quad (22)$$

Let us now verify that for a cyclic input (19) *all* allowed output configurations in a network  $U$  (21), with the above described partially distinguishable bosons, belong to the allowed set  $\mathcal{A}$  of the Fourier network  $F^{(M)}$  (18) with completely indistinguishable bosons (where there is an exponential number of such  $|\mathcal{A}| \sim N^{N-1}$ ). First, let us assume that  $k_1$  is even. The indices  $1 \leq l_1, \dots, l_N \leq M$  of output modes of  $U$ , where bosons from one of the two groups can end up, are derived from the output indices in the corresponding submatrix  $F^{(M_1)}$ , satisfying Eq. (22), by the following two rules (respectively, for the first and second blocks of  $U$ )

$$l_\alpha = 2l_\alpha^{(1)}, \quad l_{N_1+\alpha} = 2l_\alpha^{(2)} + 2M_1, \quad \alpha = 1, \dots, N_1. \quad (23)$$

Since  $N = 2N_1$  Eq. (20) is satisfied by such output mode indices

$$\sum_{\alpha=1}^N l_\alpha = (q^{(1)} + q^{(2)} + M_1)N. \quad (24)$$

Similarly, when  $k_1$  is odd we have a similar relation for the output mode indices

$$l_\alpha = 2l_\alpha^{(1)} - 1, l_{N_1+\alpha} = 2l_\alpha^{(2)} - 1 + 2M_1, \alpha = 1, \dots, N_1, \quad (25)$$

with a similar conclusion. Thus, all of the exponentially many allowed output configurations in  $U$  belong to  $\mathcal{A}$ . Though we have considered an even number of bosons, a similar example can be devised for an odd number as well.

The considered combination of a matrix  $U$  and partially distinguishable bosons is just a special simple case of many such combinations of a network and partially distinguishable bosons (see also Ref. [5]) with the allowed output configurations being an exponential subset of  $\mathcal{A}$ . For example, network  $U$  (21) relates its output modes to input modes of the same parity, but a slightly modified network consisting of  $U$  (21) followed by a “noisy” network implementing a random shift of all modes by 1 (mod  $M$ ) with probability 1/2 will have an exponential number of allowed output configurations all belonging to  $\mathcal{A}$  and no parity symmetry [24]. Due to many such possibilities, there is no way to guess, not knowing a network matrix  $U \neq F^{(M)}$ , what kind of particular feature distinguishes an exponential in  $N$  subset of  $\mathcal{A}$  which corresponds to such  $U$  and partially distinguishable bosons. An experimentalist assuming that a network matrix is close to  $F^{(M)}$  (18) and input bosons close to being completely indistinguishable, attempting a holistic assessment by the above protocol, runs only the BS regime. After a polynomial number of runs of a device, which in reality has one of the alternative possibilities of network matrix  $U \neq F^{(M)}$  and partially distinguishable bosons, the experimentalist would conclude that the input is *fully* compatible with the completely indistinguishable bosons and the matrix with  $F^{(M)}$ , since with only a polynomial number of runs one cannot tell from an experimental statistics that only some exponential in  $N$  part of  $\mathcal{A}$  is actually realized.

The conclusion is that, to close all possible loopholes in an efficient assessment protocol of BS, it is absolutely necessary to *independently* certify a network matrix, e.g., using an efficient method of Ref. [6] requiring only classical coherence. Besides eliminating the possibility of loopholes, such a certification would also guarantee that network errors are below an established acceptable level when BS is run on such a network. The theoretical basis for a network assessment is provided in Refs. [7, 8], where the effect of network matrix errors on output probability distribution of a BS device with such a network is estimated.

## II Probability of $N$ particles to gather in $\mathcal{K}$ output modes of a linear network

Let us first recall principal steps in derivation of the output probability distribution of  $N$  identical

bosons/fermions at input of a linear network  $U$  (for bosons this result was derived in Refs. [5, 9]). We will consider simultaneously both species (in case of fermions an order of creation and annihilation operators is assumed). A more general input is assumed, with  $0 \leq n_k \leq N$  particles per input mode  $k$  (fermions have linearly independent internal states in each input mode). A general input state of configuration  $\mathbf{n} = (n_1, \dots, n_M)$  reads

$$\rho(\mathbf{n}) = \sum_i q_i |\Psi_i\rangle \langle \Psi_i|, \quad (26)$$

with  $q_i \geq 0$ ,  $\sum_i q_i = 1$ , and

$$|\Psi_i\rangle = \frac{1}{\sqrt{\mu(\mathbf{n})}} \sum_{\mathbf{j}} C_{\mathbf{j}}^{(i)} \prod_{\alpha=1}^N a_{k_\alpha, j_\alpha}^\dagger |0\rangle, \quad (27)$$

where a basis state  $|j\rangle \in \mathcal{H}$  in the internal space is introduced (e.g., a basis function of spectral shape of a photon),  $\mathbf{j} = (j_1, \dots, j_N)$  and  $\mu(\mathbf{n}) = \prod_{k=1}^M n_k!$ . Permutation symmetry (anti-symmetry) of creation operators for bosons (fermions) allows to choose expansion coefficients  $C_{\mathbf{j}}^{(i)}$  symmetric (anti-symmetric) with respect to the Young subgroup  $\mathcal{S}_{\mathbf{n}} \equiv \mathcal{S}_{n_1} \otimes \dots \otimes \mathcal{S}_{n_M}$  of the symmetric group  $\mathcal{S}_N$ , where  $\mathcal{S}_{n_k}$  corresponds to permutations of internal states of particles in input mode  $k$  between themselves. Such coefficients are normalized by  $\sum_{\mathbf{j}} |C_{\mathbf{j}}^{(i)}|^2 = 1$ .

The probability of an output configuration  $\mathbf{m} = (m_1, \dots, m_M)$  is given as [5, 9, 10]

$$\hat{p}(\mathbf{m}|\mathbf{n}) = \text{Tr}(\rho(\mathbf{n})\mathcal{D}(\mathbf{m})), \quad (28)$$

where  $\rho$  is the input state Eq. (26) and  $\mathcal{D}(\mathbf{m})$  is the detection operator [5, 10] ( $|0\rangle$  is Fock vacuum state)

$$\mathcal{D}(\mathbf{m}) = \frac{1}{\mu(\mathbf{m})} \sum_{\mathbf{j}} \left[ \prod_{\alpha=1}^N b_{l_\alpha, j_\alpha}^\dagger \right] |0\rangle \langle 0| \left[ \prod_{\alpha=1}^N b_{l_\alpha, j_\alpha} \right]. \quad (29)$$

One can evaluate the trace in Eq. (28) by first expressing the input mode operators in Eq. (26) through the output ones using  $a_{k,j}^\dagger = \sum_{l=1}^M U_{k,l} b_{l,j}^\dagger$  and then employ the following identity (see also Refs. [5, 9, 10])

$$\begin{aligned} & \langle 0| \left[ \prod_{\alpha=1}^N b_{l_\alpha, j_\alpha} \right] \left[ \prod_{\alpha=1}^N b_{l'_\alpha, j'_\alpha}^\dagger \right] |0\rangle \\ &= \sum_{\sigma \in \mathcal{S}_N} \varepsilon(\sigma) \prod_{\alpha=1}^N \delta_{l'_\alpha, l_{\sigma(\alpha)}} \delta_{j'_\alpha, j_{\sigma(\alpha)}}. \end{aligned} \quad (30)$$

Substituting Eq. (26) and (29) into Eq. (28) and using Eq. (30) in the two inner products one obtains the probability of an output configuration  $\mathbf{m}$  in a linear network  $U$  in the form

$$\hat{p}(\mathbf{m}|\mathbf{n}) = \frac{1}{\mu(\mathbf{m})\mu(\mathbf{n})} \sum_{\tau, \sigma \in \mathcal{S}_N} J(\tau\sigma^{-1}) \prod_{\alpha=1}^N U_{k_{\tau(\alpha)}, l_\alpha}^* U_{k_{\sigma(\alpha)}, l_\alpha}. \quad (31)$$



where  $l_1, \dots, l_N$  are output modes,  $1 \leq l_\alpha \leq M$ , with multiplicities  $(m_1, \dots, m_M)$ , whereas function  $J(\sigma)$  and the internal state are

$$\rho^{(int)} = \sum_i q_i |\psi_i\rangle \langle \psi_i|, \quad |\psi_i\rangle \equiv \sum_{\mathbf{j}} C_{\mathbf{j}}^{(i)} \prod_{\alpha=1}^N \otimes |j_\alpha\rangle \quad (32)$$

and

$$J(\sigma) = \varepsilon(\sigma) \text{Tr}(\rho^{(int)} P_\sigma), \quad \varepsilon(\sigma) = \begin{cases} 1, & \text{Bosons,} \\ \text{sgn}(\sigma), & \text{Fermions,} \end{cases} \quad (33)$$

where  $P_\sigma \prod_{\alpha=1}^N \otimes |j_\alpha\rangle = \prod_{\alpha=1}^N \otimes |j_{\sigma^{-1}(\alpha)}\rangle$  is the operator representation of  $\sigma$  in  $\mathcal{H}^{\otimes N}$ . Note that the permutation symmetry (anti-symmetry) of the input state (26)-(27) for bosons (fermions), i.e.,  $P_\pi \rho^{(int)} = \rho^{(int)} P_\pi = \varepsilon(\pi) \rho^{(int)}$  for any  $\pi \in \mathcal{S}_n$ , implies that

$$J(\sigma\pi) = J(\pi\sigma) = J(\sigma), \quad \forall \pi \in \mathcal{S}_n. \quad (34)$$

A network input consists of *completely indistinguishable* bosons (fermions) if the corresponding  $J$ -function reads  $J^{(id)}(\sigma) = \varepsilon(\sigma)$  [5, 10] (in case of fermions  $\mu(\mathbf{n}) = 1$ ). This case allows one to completely neglect the internal degrees of freedom. Probabilities at a network output are expressed through the usual matrix permanent and determinant, respectively. The simplest case of completely indistinguishable particles consists of all particles being in the same internal state  $|\phi\rangle$ , giving  $\rho^{(int)} = (|\phi\rangle \langle \phi|)^{\otimes N}$ .

The other limiting case, which may be identified as the *classical case*, since the output probabilities are the same as in the case of classical particles, corresponds to a “block-structured”  $J$ -function (see also Refs. [5, 9])

$$J^{(d)}(\sigma) = \sum_{\pi \in \mathcal{S}_n} \delta_{\sigma, \pi}. \quad (35)$$

Function  $J(\sigma)$  of Eq. (35) appears when the internal states of identical particles from different input modes become orthogonal:  $\text{Tr}\{\rho^{(int)} P_\sigma\} = 0$  for  $\sigma \notin \mathcal{S}_n$ , whereas (by the symmetry of  $C_{\mathbf{j}}$ ) we always have  $\varepsilon(\sigma) \text{Tr}\{\rho^{(int)} P_\sigma\} = 1$  for  $\sigma \in \mathcal{S}_n$  (i.e., distinguishable particles from the same input mode cannot be discriminated by a linear network from the completely indistinguishable bosons). Note that the subgroup  $\mathcal{S}_n$  acts as identity on the indices  $k_1, \dots, k_N$  of matrix  $U$  in Eq. (31), thus the sum over  $\mathcal{S}_n$  in Eq. (35) cancels  $\mu(\mathbf{n})$  in the denominator in Eq. (31), resulting in the familiar formula for the probability in the classical case, expressed through the matrix permanent of doubly stochastic matrix with elements  $|U_{kl}|^2$ . Obviously, for single particles at input ( $n_k \leq 1$ ) we have  $J^{(d)}(\sigma) = \delta_{\sigma, I}$ .

The total probability of detecting all  $N$  input particles at a preselected (and fixed) set of  $1 \leq \mathcal{K} \leq M$  output modes, say, the first  $\mathcal{K}$  modes, is a sum of  $\hat{p}(\mathbf{m}|\mathbf{n})$  (31)

with  $m_{\mathcal{K}+1} = \dots m_M = 0$ . We have

$$\begin{aligned} p_N(J) &= \sum_{\mathbf{m}}' \hat{p}(\mathbf{m}|\mathbf{n}) = \frac{1}{N!} \sum_{l_1=1}^{\mathcal{K}} \dots \sum_{l_N=1}^{\mathcal{K}} \\ &\times \frac{1}{\mu(\mathbf{n})} \sum_{\tau, \sigma \in \mathcal{S}_N} J(\tau\sigma^{-1}) \prod_{\alpha=1}^N U_{k_{\tau(\alpha)}, l_\alpha}^* U_{k_{\sigma(\alpha)}, l_\alpha} \\ &= \frac{1}{\mu(\mathbf{n})} \sum_{\sigma' \in \mathcal{S}_N} J(\sigma') \prod_{\alpha=1}^N H_{\alpha, \sigma'(\alpha)} = \frac{1}{\mu(\mathbf{n})} \sum_{\sigma \in \mathcal{S}_N} J(\sigma) \Pi_{I, \sigma} \\ &= \frac{1}{\mu(\mathbf{n})} \sum_{\tau, \tau' \in \mathcal{S}_N} \theta^*(\tau) \theta(\tau') \Pi_{\tau, \tau'}. \end{aligned} \quad (36)$$

where we have transformed the sum over output configurations  $\mathbf{m}$  into that over output mode indices  $l_1, \dots, l_N$  with the combinatorial coefficient  $\mu(\mathbf{m})/N!$ , defined the following matrices

$$H_{\alpha, \beta} \equiv \sum_{l=1}^{\mathcal{K}} U_{k_\alpha, l} U_{k_\beta, l}^*, \quad \Pi_{\sigma, \tau}(H) = \prod_{\alpha=1}^N H_{\sigma(\alpha), \tau(\alpha)}, \quad (37)$$

reordered the product  $\prod_{\alpha=1}^N H_{\sigma(\alpha), \tau(\alpha)} = \prod_{\alpha=1}^N H_{\alpha, \tau\sigma^{-1}(\alpha)}$ , and defined  $\sigma' = \tau\sigma^{-1}$  (the sum over  $\tau$  cancels the factor  $1/N!$ ).

Note that for a p.s.d. Hermitian  $H$ , as a principal submatrix of the tensor product matrix  $H^{\otimes N}$  the Schur power matrix is a p.s.d. Hermitian as well.

Consider now the limit cases. In case of the completely indistinguishable particles  $\theta^{(id)}(\sigma) = \frac{\varepsilon(\sigma)}{\sqrt{N!}}$  (since  $J^{(id)}(\sigma) = \varepsilon(\sigma)$ ). Eq. (36) gives for bosons and fermions:

$$p_N^{(B)} = \frac{\text{per}(H)}{\mu(\mathbf{n})}, \quad p_N^{(F)} = \det(H) \delta_{\mu(\mathbf{n}), 1}, \quad (38)$$

where  $\text{per}(A) \equiv \sum_{\sigma} \prod_{\alpha=1}^N A_{\alpha, \sigma(\alpha)}$ . For classical particles (e.g., distinguishable bosons or fermions from different input modes) we get from Eq. (35) that

$$\theta^{(d)}(\sigma) = \frac{1}{\sqrt{\mu(\mathbf{n})}} \sum_{\pi \in \mathcal{S}_n} \delta_{\sigma, \pi}, \quad (39)$$

which results in

$$p_N^{(d)} = \prod_{\alpha=1}^N H_{\alpha, \alpha}. \quad (40)$$

The following order of the limit-case probabilities can be easily established

$$p_N^{(F)} \leq p_N^{(d)} \leq p_N^{(B)}, \quad (41)$$

valid for *arbitrary* number of particles per input mode. Indeed, for a p.s.d. Hermitian matrix  $H$ , which we rearrange in a block-matrix form

$$H = \begin{pmatrix} H^{(1,1)} & H^{(1,2)} \\ H^{(2,1)} & H^{(2,2)} \end{pmatrix}. \quad (42)$$



the following inequality is known for the matrix determinant [11]:

$$\det(H) \leq \det(H^{(1,1)})\det(H^{(2,2)}). \quad (43)$$

Similarly, for the matrix permanent [12]

$$\text{per}(H) \geq \text{per}(H^{(1,1)})\text{per}(H^{(2,2)}). \quad (44)$$

By repeated application of Eqs. (43) and (44) it is easy to demonstrate that Eq. (41) holds.

### III Factorization of $J$ -function and its representation through a density matrix

To show that the probability  $p_N^{(B)}$  ( $p_N^{(F)}$ ) for the completely indistinguishable bosons (fermions) corresponds to the absolute maximum (minimum) over *arbitrary* input states of particles in a given configuration  $\mathbf{n}$ , one has to know to what class of functions the physical  $J$ -functions, i.e., describing an input of a linear network, belong. Let us show that any normalized by  $J(I) = 1$  p.s.d. function  $J(\sigma)$  can be represented in the form of Eq. (33) with some state  $\rho^{(int)}$  (32). Since  $\text{sgn}(\sigma)J(\sigma)$  is also a normalized p.s.d function, it is sufficient to consider bosons,  $\varepsilon(\sigma) = 1$ .

Consider a linear subspace  $\mathcal{L} \subset \mathcal{H}^{\otimes N}$  defined as the linear span of vectors  $|\sigma\rangle \equiv P_\sigma|I\rangle$  for  $\sigma \in \mathcal{S}_N$ , where  $|I\rangle \equiv |\phi_1\rangle \otimes \dots \otimes |\phi_N\rangle$  for some arbitrary orthonormal vectors  $|\phi_k\rangle \in \mathcal{H}$ ,  $\langle\phi_k|\phi_l\rangle = \delta_{k,l}$  (note that  $\langle\sigma|\tau\rangle = \delta_{\sigma,\tau}$  and  $P_\pi = \sum_\tau |\pi\tau\rangle\langle\tau|$  when restricted to  $\mathcal{L}$ ). Using that  $P_\sigma|\tau\rangle = |\sigma\tau\rangle$  and  $\langle\pi|P_\sigma|\pi\rangle = \delta_{\sigma,I}$ , by starting from a trivial identity we get

$$\begin{aligned} J(\sigma) &= \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \langle\pi| \left[ \sum_{\tau} J(\tau) P_\tau^\dagger \right] P_\sigma |\pi\rangle \\ &= \text{Tr} \left\{ \rho^{(int)} P_\sigma \right\}, \end{aligned} \quad (45)$$

where the trace is taken in  $\mathcal{H}^{\otimes N}$  and we have introduced a p.s.d. Hermitian operator (density matrix)  $\rho^{(int)}$  in the Hilbert space  $\mathcal{H}^{\otimes N}$

$$\rho^{(int)} \equiv \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} J(\tau) \sum_{\pi \in \mathcal{S}_N} |\pi\rangle\langle\tau\pi|. \quad (46)$$

Obviously,  $\text{Tr}\{\rho^{(int)}\} = J(I) = 1$ . Positivity of  $\rho^{(int)}$  would follow from the explicit form (now for bosons and fermions)

$$\begin{aligned} \rho^{(int)} &= \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} |\Phi_\tau\rangle\langle\Phi_\tau|, \\ |\Phi_\tau\rangle &\equiv \sum_{\sigma \in \mathcal{S}_N} \xi(\sigma\tau) P_\sigma \{|\phi_1\rangle \otimes \dots \otimes |\phi_N\rangle\}, \end{aligned} \quad (47)$$

where  $\xi(\sigma) \equiv \varepsilon(\sigma)\theta^*(\sigma)$  for the factorizing function  $\theta(\sigma)$  (see Eq. (51) below).

Let us sketch the prove of factorization of a p.s.d. Hermitian function  $J(\sigma)$  (see Ref. [13]). Consider an operator  $\mathcal{J}$  in  $\mathcal{L}$

$$\mathcal{J} \equiv \sum_{\sigma \in \mathcal{S}_N} J(\sigma) P_\sigma^\dagger, \quad (48)$$

given by following matrix

$$\mathcal{J}_{\nu,\tau} = \langle\nu|\mathcal{J}|\tau\rangle = J(\nu^{-1}\tau), \quad (49)$$

which, by assumption that  $J(\sigma)$  is a p.s.d. Hermitian function, is a p.s.d. Hermitian matrix. Since operators having the form given by Eq. (48) constitute an sub-algebra of operators in  $\mathcal{L}$ , the p.s.d. Hermitian operator  $\mathcal{J}$  (48) can be factorized by an operator  $\mathcal{B} \in \mathcal{L}$

$$\mathcal{J} = \mathcal{B}^\dagger \mathcal{B}, \quad \mathcal{B} = \sum_{\sigma} \theta(\sigma) P_\sigma^\dagger. \quad (50)$$

By the group rule  $P_\sigma P_\tau = P_{\sigma\tau}$  Eq. (50), in the matrix form, is equivalent to the factorization

$$J(\sigma) = \sum_{\tau \in \mathcal{S}_N} \theta^*(\tau) \theta(\tau\sigma), \quad \sum_{\sigma \in \mathcal{S}_N} |\theta(\sigma)|^2 = 1. \quad (51)$$

The symmetry (34) means that the factorizing function can be chosen to satisfy

$$\theta(\sigma\pi) = \theta(\sigma), \quad \forall \pi \in \mathcal{S}_n. \quad (52)$$

(In contrast,  $\theta(\tau\sigma)$  for all  $\tau \in \mathcal{S}_N$  is the same factorization with a different order of terms).

### IV Derivation of the average probability formulae

Though the average quantum detection probabilities follow from a simple symmetry argument for a Haar-random unitary  $U$ , the classical case requires a bit more of insight. Here these results are derived by direct evaluation demonstrating also the validity of the classical formula for  $M \gg 1$  (arbitrary  $\mathcal{K}$  and  $N$ ), as observed in numerical simulations. The following identity will be employed

$$\begin{aligned} &\left\langle \prod_{\alpha=1}^N U_{k_\alpha, l_\alpha} U_{k'_\alpha, l'_\alpha}^* \right\rangle \\ &= \sum_{\nu, \tau \in \mathcal{S}_N} \mathcal{W}(M, \nu\tau^{-1}) \prod_{\alpha=1}^N \delta_{k'_\alpha, k_{\nu(\alpha)}} \delta_{l'_\alpha, l_{\tau(\alpha)}}, \end{aligned} \quad (53)$$

where  $\mathcal{W}(M, \sigma)$  is the Weingarten function of the unitary group [14, 16] which depends only on the cycle structure of the relative permutation  $\sigma = \nu\tau^{-1}$ , i.e., the sequence of numbers  $(c_1(\sigma), \dots, c_N(\sigma))$  of cycles of lengths  $(1, \dots, N)$  in its cycle decomposition (for more details see Ref. [17]). By application of Eq. (53) we have

$$\left\langle \prod_{\alpha=1}^N U_{k_\alpha, l_\alpha} U_{k_{\sigma(\alpha)}, l_\alpha}^* \right\rangle = \sum_{\nu \in \mathcal{S}_n} \sum_{\tau \in \mathcal{S}_m} \mathcal{W}(M, \nu\sigma\tau) \quad (54)$$

with summation over permutation invariance subgroups  $\mathcal{S}_{\mathbf{n}}$  and  $\mathcal{S}_{\mathbf{m}}$  of the input and output indices, respectively. Then, from Eq. (31) we have for the average over Haar-random network (here the output indices vary between 1 and  $\mathcal{K}$  and, since we sum over the output modes  $l_1, \dots, l_N$ , one of permutations in  $J$  is redundant, giving a  $N!$  factor)

$$\begin{aligned} \langle p_N(J) \rangle &= \frac{N!}{\mu(\mathbf{n})} \sum'_{\mathbf{m}} \frac{1}{\mu(\mathbf{m})} \sum_{\sigma \in \mathcal{S}_N} J(\sigma) \langle \prod_{\alpha=1}^N U_{k_\alpha, l_\alpha} U_{k_{\sigma(\alpha)}, l_\alpha}^* \rangle \\ &= \sum_{\sigma \in \mathcal{S}_N} J(\sigma) \sum'_{\mathbf{m}} \frac{N!}{\mu(\mathbf{m})} \sum_{\tau \in \mathcal{S}_{\mathbf{m}}} \mathcal{W}(M, \sigma\tau), \end{aligned} \quad (55)$$

where  $\sum'$  is over all occupations  $\mathbf{m}$  of  $\mathcal{K}$  output modes and we have used the property of  $J(\sigma)$  in Eq. (34) and that the number of  $\nu \in \mathcal{S}_{\mathbf{n}}$  is  $|\mathcal{S}_{\mathbf{n}}| = \mu(\mathbf{n})$ . Let us now consider separately bosons, fermions and distinguishable particles for a general input  $\mu(\mathbf{n}) \geq 1$ .

In case of the completely indistinguishable bosons,  $J(\sigma) = 1$ , we obtain from Eqs. (55) and (54)

$$\begin{aligned} \langle p_N^{(B)} \rangle &= \sum'_{\mathbf{m}} \frac{N!}{\mu(\mathbf{m})} \sum_{\sigma \in \mathcal{S}_N} \sum_{\tau \in \mathcal{S}_{\mathbf{m}}} \mathcal{W}(M, \sigma\tau) \\ &= \frac{(\mathcal{K} + N - 1)!}{(\mathcal{K} - 1)!} \sum_{\sigma \in \mathcal{S}_N} \mathcal{W}(M, \sigma), \end{aligned} \quad (56)$$

where we have used that  $\mathcal{S}_{\mathbf{m}} \subset \mathcal{S}_N$  and  $|\mathcal{S}_{\mathbf{m}}| = \mu(\mathbf{m})$  (the first sum gives the number of Fock states of  $N$  bosons in  $\mathcal{K}$  modes scaled by  $N!$ ). Observing that for  $\mathcal{K} = M$  the probability must be equal to 1, we get the sum of  $\mathcal{W}$ -functions in Eq. (56)

$$\sum_{\sigma \in \mathcal{S}_N} \mathcal{W}(M, \sigma) = \frac{(M - 1)!}{(M + N - 1)!}. \quad (57)$$

Combining Eqs. (56) and (57) we get the final expression for  $\langle p_N^{(B)} \rangle$ .

In case of the completely indistinguishable fermions  $J(\sigma) = \text{sgn}(\sigma)$  and  $\mu(\mathbf{n}) = 1$  (no two or more particles per mode). We have from Eqs. (55) and (54)

$$\begin{aligned} \langle p_N^{(F)} \rangle &= N! \sum'_{\mathbf{m}} \sum_{\sigma \in \mathcal{S}_N} \text{sgn}(\sigma) \mathcal{W}(M, \sigma), \\ &= \frac{\mathcal{K}!}{(\mathcal{K} - N)!} \sum_{\sigma \in \mathcal{S}_N} \text{sgn}(\sigma) \mathcal{W}(M, \sigma) \end{aligned} \quad (58)$$

(the first sum is the number of Fock states of  $N$  fermions in  $\mathcal{K}$  modes). Setting  $\mathcal{K} = M$  in Eq. (58) we obtain

$$\sum_{\sigma \in \mathcal{S}_N} \text{sgn}(\sigma) \mathcal{W}(M, \sigma) = \frac{(M - N)!}{M!}. \quad (59)$$

Eqs. (58) and (59) result in the final expression for  $\langle p_N^{(F)} \rangle$ .

In case of distinguishable particles from Eqs. (35) and (55) we get

$$\begin{aligned} \langle p_N^{(d)} \rangle &= \sum_{l_1=1}^{\mathcal{K}} \dots \sum_{l_N=1}^{\mathcal{K}} \sum_{\nu \in \mathcal{S}_{\mathbf{n}}} \sum_{\tau \in \mathcal{S}_{\mathbf{m}}} \mathcal{W}(M, \nu\tau) \\ &= \sum_{\nu \in \mathcal{S}_{\mathbf{n}}} \sum_{\sigma \in \mathcal{S}_N} \mathcal{W}(M, \nu\sigma) \sum_{l_1=1}^{\mathcal{K}} \dots \sum_{l_N=1}^{\mathcal{K}} \prod_{\alpha=1}^N \delta_{l_\alpha, l_{\sigma(\alpha)}} \\ &= \sum_{\nu \in \mathcal{S}_{\mathbf{n}}} \sum_{\sigma \in \mathcal{S}_N} \mathcal{W}(M, \nu\sigma) \mathcal{K}^{\#\sigma} \end{aligned} \quad (60)$$

where we have set  $\#\sigma \equiv c_1(\sigma) + \dots + c_N(\sigma)$  (the total number of cycles in the cycle decomposition of  $\sigma$ ) and used that  $|\mathcal{S}_{\mathbf{n}}| = \mu(\mathbf{n})$  and  $\prod_{\alpha=1}^N \delta_{l_\alpha, l_{\sigma(\alpha)}} = \sum_{\tau \in \mathcal{S}_{\mathbf{m}}} \delta_{\sigma, \tau}$ . Eq. (60) must coincide with Eq. (56) for all particles in the same input mode,  $\mu(\mathbf{n}) = N!$ , since the limit of distinguishable particles Eq. (35) is obtained by making identical particles from different modes distinguishable, while particles from the same input mode behave as the completely indistinguishable bosons. On the other hand, for a single-particle input it has a form quite different from that of Eq. (56) for a similar input of the completely indistinguishable bosons. Apparently, the general expression would have quite a cumbersome form. Consider the special case of single particles. Using an asymptotic form of  $\mathcal{W}$  for  $M \gg 1$  [16]

$$\begin{aligned} \mathcal{W}(M, \sigma) &= \frac{(-1)^N}{M^{2N}} \prod_{s=1}^N (-M g_s)^{c_s(\sigma)} \left( 1 + O\left(\frac{1}{M^2}\right) \right) \\ g_s &= \frac{(2s - 2)!}{s!(s - 1)!}, \end{aligned} \quad (61)$$

we obtain from Eq. (60) the leading term in the average classical probability as a cycle sum

$$\langle p_N^{(d)} \rangle \approx \frac{(-1)^N}{M^{2N}} \sum_{\sigma \in \mathcal{S}_N} \prod_{s=1}^N (-\mathcal{K} M g_s)^{c_s(\sigma)} = \frac{(-1)^N N!}{M^{2N}} Z_N, \quad (62)$$

with  $(t_s = -\mathcal{K} M g_s)$

$$Z_N \equiv \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \prod_{s=1}^N t_s^{c_s(\sigma)}. \quad (63)$$

The cycle sum is evaluated by the generating function method (see, for instance, Ref. [17]) which satisfies the following identity

$$\mathcal{F}(x) \equiv \sum_{N \geq 1} Z_N x^N = \exp \left\{ \sum_{s \geq 1} t_s \frac{x^s}{s} \right\}. \quad (64)$$

In our case (after evaluation of a table sum [18]) we get  $\mathcal{F}(x)$  to be

$$\begin{aligned} \mathcal{F}(x) &= \exp \left\{ -\mathcal{K} M \sum_{s=1}^{\infty} \frac{(2s - 2)!}{(s!)^2} x^s \right\} \\ &= \left( \frac{2}{1 + \sqrt{1 - 4x}} \exp \{ -[1 - \sqrt{1 - 4x}] \} \right)^{\mathcal{K} M}. \end{aligned} \quad (65)$$

For  $\mathcal{KM} \gg 1$ , the leading order reads

$$Z_N = \frac{1}{N!} \frac{d^N \mathcal{F}(0)}{dx^N} \approx \frac{(-\mathcal{KM})^N}{N!}, \quad (66)$$

due to the fact that the derivative of  $\mathcal{F}(x)$ ,

$$\frac{d\mathcal{F}(x)}{dx} = \frac{-2\mathcal{KM}}{\sqrt{1-4x}(1+\sqrt{1-4x})} \mathcal{F}(x), \quad (67)$$

has as a factor a slowly varying function of  $|x| \ll 1$  (in comparison with  $\mathcal{F}(x)$ ). Hence  $\frac{d^N \mathcal{F}(0)}{dx^N} = (-\mathcal{KM})^N [1 - N(N-1)/2\mathcal{KM} + \dots]$ . Substituting Eq. (66) into Eq. (62) we obtain  $\langle p_N^{(d)} \rangle$  up to a factor  $(1 + O(\frac{N^2}{\mathcal{KM}}))$ .

## V Classical computations for the assessment protocol are only polynomial in $N$

The assessment protocol for the standard version of BS is dependent on  $p_N^{(B)} = \text{per}(H)$ , where  $H$  is a p.s.d. Hermitian matrix built from a submatrix of a network matrix  $U$  according to Eq. (6) of the main text. Such a matrix  $H$  can be rewritten as

$$H_{\alpha,\beta} = \delta_{\alpha,\beta} - \sum_{l=\mathcal{K}+1}^M U_{k_\alpha,l} U_{k_\beta,l}^* \equiv (I - \Phi)_{\alpha,\beta}, \quad (68)$$

where  $k_1, \dots, k_N$  are input indices. Below we show that, in the dilute limit,  $NL/M \rightarrow 0$  as  $N \rightarrow \infty$ , with a high probability over a choice of  $L = M - \mathcal{K}$  output modes, only polynomial in  $N$  computations are required to estimate  $\text{per}(H)$  of Eq. (68) to an arbitrary error  $\epsilon$  polynomial in  $N^{-1}$  [25]. Since we can select  $L$  output modes completely arbitrary, one can always find a suitable set in a polynomial number of trials. Moreover, an approximation formula to the required polynomial accuracy is given (Eq. (73) below).

We will use a general formula for the matrix permanent of a sum of two matrices [19], in our case

$$\text{per}(H) = 1 + \sum_{r=1}^N (-1)^r \sum_{\alpha_1 < \dots < \alpha_r} \text{per}(\Phi[\vec{\alpha}, \vec{\alpha}]), \quad (69)$$

where  $\Phi[\vec{\alpha}, \vec{\alpha}]$  is an  $r$ -dimensional submatrix of  $\Phi$  built on the rows and columns  $\vec{\alpha} = (\alpha_1, \dots, \alpha_r)$ . Now, since for  $NL/M \ll 1$  the result on the r.h.s. of Eq. (69) is close to 1 (see Eq. (12) of the main text), we expect that only a few terms in this expression are required for estimating  $\text{per}(H)$  to a polynomial in  $N^{-1}$  error.

Indeed, let us consider how many terms on the r.h.s. of Eq. (69) are required on average over the Haar-random networks. By application of the general result given by Eq. (12) in the main text (see also the previous section) we have

$$\langle \text{per}(\Phi[\vec{\alpha}, \vec{\alpha}]) \rangle = \frac{L(L+1) \cdot \dots \cdot (L+r-1)}{M(M+1) \cdot \dots \cdot (M+r-1)}, \quad (70)$$

hence, for  $r \ll \sqrt{N}$  and  $r \ll \sqrt{L}$  we obtain

$$\langle T_r \rangle \equiv \sum_{\alpha_1 < \dots < \alpha_r} \langle \text{per}(\Phi[\vec{\alpha}, \vec{\alpha}]) \rangle = (1 + \varepsilon) \frac{1}{r!} \left( \frac{NL}{M} \right)^r, \quad (71)$$

where  $\varepsilon \sim r^2[1/N + 1/L] \ll 1$ .

To estimate  $\text{per}(H)$  to an error  $\epsilon = O(N^{-\kappa})$ , for  $\kappa > 0$ , we can truncate the sum on the r.h.s. of Eq. (69) at an order  $s$  satisfying  $\langle T_s \rangle \ll \epsilon$ . To have a single scale  $N^{-1}$ , let us concentrate on the case given by  $M = bN^{2+\delta}$  for  $b = O(1)$  and  $\delta > 0$  and  $L = aN$  with  $a = O(1)$ . In this case, averaging over Haar-random  $U$  gives  $\langle \text{per}(H) \rangle = 1 - O(N^{-\delta})$ . The truncation order  $s$  must satisfy

$$\langle T_s \rangle \approx \frac{1}{s!} \left( \frac{a}{bN^\delta} \right)^s \ll \epsilon = O(N^{-\kappa}), \quad (72)$$

i.e., for  $N \rightarrow \infty$  one must choose  $s > \kappa/\delta$ . Let us set  $s = (\kappa + 1)/\delta$  (note that  $s$  is an independent of  $N$  number) then for  $N \gg 1$  the l.h.s. of Eq. (72) is smaller by a factor  $O(N^{-1})$  than the r.h.s.. Now, let us show that the number of required computations (flops)  $\mathcal{C}_N$  in the summation of the r.h.s. of Eq. (69) to the above truncation order  $s = O(1)$  scales only polynomially in  $N$  as  $N \rightarrow \infty$ . Since computation of a matrix permanent of dimension  $r$  requires  $r2^r$  operations (flops) by Ryser's algorithm [19], we can estimate the total number of flops as follows. First we note that Eq. (69) truncated at an order  $s$  has an equivalent simpler expression (to the same accuracy)

$$\text{per}(H) \approx 1 + \sum_{\alpha_1 < \dots < \alpha_s} \{ \text{per}(\Phi[\vec{\alpha}, \vec{\alpha}]) - 1 \} \quad (73)$$

(we have taken into account that indices  $\alpha_1 < \dots < \alpha_r$  with  $r \leq s$  are contained in one of  $\alpha_1 < \dots < \alpha_s$  and used Eq. (69) to sum the terms for each subset  $\alpha_1 < \dots < \alpha_s$ ). From Eq. (73) we obtain

$$\mathcal{C}_N \sim \frac{N!}{(N-s)!s!} s2^s = O\left(N^{\frac{\kappa+1}{\delta}}\right). \quad (74)$$

where we have used  $s = (\kappa + 1)/\delta$  for the approximation error  $\epsilon = O(N^{-\kappa})$ .

The estimates in Eqs. (74) and (73) (and that of Eq. (71)) have a high probability, since by Chebyshev's inequality (for  $N \gg 1$ )

$$\text{Pr}(T_s > \epsilon) \leq \frac{\langle T_s \rangle}{\epsilon} = O(N^{-1}) \ll 1, \quad (75)$$

where we have used Eq. (72) [26]. Since we select  $L = M - \mathcal{K}$  output modes for the protocol completely arbitrary from the output modes of a network, if it turns out that for our choice of  $L$  modes for a particular network  $U$  we need more than  $\mathcal{C}_N$  flops, we can select again (e.g., by using the Gaussian approximation [20], the probability that no suitable choice can be found among  $O(N)$  non-intersecting subsets of size  $L = O(N)$  from the total of  $M = O(N^{2+\delta})$  modes decreases to zero at least as  $O(N^{-N})$ .

One important example is the test against the distinguishable particles, for which one can set  $\kappa = 2(1 + \delta)$ . Indeed, the difference  $\langle p_N^{(B)} - p_N^{(d)} \rangle$  between the full quantum and the classical cases (see the main text) reads in our case

$$\langle p_N^{(B)} - p_N^{(d)} \rangle = O\left(\frac{LN^2}{\kappa M}\right) = O(N^{-1-2\delta}), \quad (76)$$

therefore, to distinguish the two cases requires using only polynomial computations by the above analysis.

## VI Effect of the distinguishability error on probability $p_N$ and its Haar-average

Consider an input state consisting of single bosons from  $N$  sources (we will consider uncorrelated sources, the case of classically correlated ones is a trivial extension). The corresponding internal state reads

$$\rho^{(int)} = \rho^{(1)} \otimes \dots \otimes \rho^{(N)}, \quad (77)$$

where  $\rho^{(\alpha)}$  is an internal state of the  $\alpha$ th boson. The ideal bosonic input corresponds (in the case of only classically correlated inputs) to all  $\rho^{(\alpha)} = |\phi\rangle\langle\phi|$ . Hence, let us assume that each source produces a boson in a state close to a pure state

$$\rho^{(\alpha)} = |\phi\rangle\langle\phi| - \delta\rho^{(\alpha)} \quad (78)$$

with the indistinguishability fidelity defined as  $F_\alpha \equiv \langle\phi|\rho^{(\alpha)}|\phi\rangle$ . To derive the scaling law for small variations  $1 - F_\alpha = \langle\phi|\delta\rho^{(\alpha)}|\phi\rangle \ll 1$  we will assume that sources emit bosons with a mean fidelity  $F$ . Keeping only the first order term in  $1 - F$  we obtain from Eq. (77)

$$\begin{aligned} \rho^{(int)} &\approx (|\phi\rangle\langle\phi|)^{\otimes N} \\ &- \sum_{\alpha=1}^N (|\phi\rangle\langle\phi|)^{\otimes(\alpha-1)} \otimes \delta\rho^{(\alpha)} \otimes (|\phi\rangle\langle\phi|)^{\otimes(N-\alpha)}. \end{aligned} \quad (79)$$

Taking into account that  $\text{Tr}\{\delta\rho^{(\alpha)}\} = 0$ , we obtain from the definition (33) and Eq. (79) up to the first order in  $1 - F$

$$J(\sigma) = \text{Tr}\{\rho^{(int)} P_\sigma\} \approx 1 - (1 - F)[N - c_1(\sigma)], \quad (80)$$

where  $c_1(\sigma)$  is the number of 1-cycles (fixed points) in the cycle decomposition of permutation  $\sigma$  ( $c_1(\sigma)$  terms in the sum in Eq. (79) give zero).

Let us first derive the scaling law for the standard version of BS. Substitution of Eq. (80) into Eq. (36) gives (with  $p_N^{(B)} = \text{per}(H)$ )

$$p_N(J) \approx p_N^{(B)} - (1 - F) \left[ N p_N^{(B)} - \sum_{\sigma \in \mathcal{S}_N} c_1(\sigma) \prod_{\alpha=1}^N H_{\alpha, \sigma(\alpha)} \right]. \quad (81)$$

The second term in the brackets on the r.h.s. of Eq. (81) is different from the plain matrix permanent by a factor at each term, counting how many diagonal elements of matrix  $H$  it contains. It can be represented in a computable form by observing that the same counting is done by multiplying diagonal elements of  $H$  by a dummy variable  $x$  and application of a derivative w.r.t.  $x$  at  $x = 1$ . Thus we have derived

$$p_N(J) \approx p_N^{(B)} - (1 - F) \left[ N p_N^{(B)} - \frac{d}{dx} \text{per}\{H(x)\}_{x=1} \right] \quad (82)$$

with

$$H(x) \equiv H + (x - 1) \text{diag}(H_{11}, \dots, H_{NN}) \quad (83)$$

The derivative can be evaluated from the Lagrange representation of the polynomial  $\text{per}\{H(x)\}$ , thus it requires approximating  $N + 1$  matrix permanents of the p.s.d. Hermitian matrix  $H(x)$  at distinct values of  $x$  to some small error  $\epsilon$ , which can be chosen inversely polynomial in  $N$ . In the previous section it is shown that this requires only polynomial in  $N$  computations.

Now, let us derive a variant of the scaling law which is specifically tailored for the scattershot BS. Assuming that  $M \gg N^2$  (so that one can use the Gaussian approximation for the elements of the Haar-random unitary matrix  $U$  [20]) we use the expression (80) into the average probability of detecting all  $N$  photons in  $\mathcal{K}$  output modes (i.e.,  $\mathbf{m} = (m_1, \dots, m_{\mathcal{K}}, 0, \dots, 0)$ , see also Eqs. (31) and (36)) and recall the formula for the average probability  $\langle p_N^{(B)} \rangle$  (in the dilute limit  $M \gg N^2$ ):

$$\begin{aligned} \langle p_N(J) \rangle &= \sum_{\mathbf{m}}' \frac{N!}{\mu(\mathbf{m})} \sum_{\sigma \in \mathcal{S}_N} J(\sigma) \langle \prod_{\alpha=1}^N U_{k_{\sigma(\alpha)}, l_\alpha}^* U_{k_\alpha, l_\alpha} \rangle \\ &\approx \langle p_N^{(B)} \rangle - (1 - F) \frac{N!}{M^N} \sum_{\mathbf{m}}' \frac{1}{\mu(\mathbf{m})} \sum_{\sigma \in \mathcal{S}_{\mathbf{m}}} [N - c_1(\sigma)] \\ &= \langle p_N^{(B)} \rangle - (1 - F) \frac{N(N-1)}{M} \langle p_{N-1}^{(B)} \rangle \end{aligned} \quad (84)$$

(note that the final result has an appealing form, apparently indicating that it might be valid without assuming the dilute limit  $M \gg N^2$ ). Mathematical details in the derivation of Eq. (84) are as follows. We have used an approximate independence of  $2N$  matrix elements of  $U$  for  $M \gg N^2$  and the Gaussian approximation [20] in computing the average of the product, with  $\langle |U_{kl}|^2 \rangle \approx 1/M$  and for single bosons at input ( $k_\alpha \neq k_\beta$  for  $\alpha \neq \beta$ )

$$\langle \prod_{\alpha=1}^N U_{k_{\sigma(\alpha)}, l_\alpha}^* U_{k_\alpha, l_\alpha} \rangle \approx \frac{1}{M^N} \sum_{\pi \in \mathcal{S}_{\mathbf{m}}} \delta_{\sigma, \pi}, \quad (85)$$

since all permutations of the output indices belonging to the output subgroup  $\mathcal{S}_{\mathbf{m}}$  equally contribute to the result. The double sum in the line before the last in Eq. (84) (over all  $\mathbf{m}$  and the subgroup  $\mathcal{S}_{\mathbf{m}}$ )

$$R = \sum_{\mathbf{m}}' \frac{1}{\mu(\mathbf{m})} \sum_{\sigma \in \mathcal{S}_{\mathbf{m}}} [N - c_1(\sigma)] \quad (86)$$

is computed as follows. We have a sum over permutations

$$\begin{aligned} \frac{1}{\mu(\mathbf{m})} \sum_{\sigma \in \mathcal{S}_{\mathbf{m}}} c_1(\sigma) &= \frac{1}{\mu(\mathbf{m})} \sum_{\sigma_1 \in \mathcal{S}_{m_1}} \cdots \sum_{\sigma_{\mathcal{K}} \in \mathcal{S}_{m_{\mathcal{K}}}} \sum_{l=1}^{\mathcal{K}} c_1(\sigma_l) \\ &= \mathcal{K} - \sum_{l=1}^{\mathcal{K}} \delta_{m_l, 0}, \end{aligned} \quad (87)$$

i.e., the average number of fixed points (over all permutations) in each permutation group  $\mathcal{S}_{m_l}$  is equal to 1 [17], whereas if  $m_l = 0$  there is no corresponding contribution. The first sum in Eq. (86) is then easily computed

$$\begin{aligned} R &= \sum_{\mathbf{m}}' \left( N - \mathcal{K} + \sum_{l=1}^{\mathcal{K}} \delta_{m_l, 0} \right) \\ &= (N - \mathcal{K}) \frac{(\mathcal{K} + N - 1)!}{N!(\mathcal{K} - 1)!} + \mathcal{K} \frac{(\mathcal{K} + N - 2)!}{N!(\mathcal{K} - 2)!} \\ &= (N - 1) \frac{(\mathcal{K} + N - 2)!}{(N - 1)!(\mathcal{K} - 1)!}, \end{aligned} \quad (88)$$

where the summation gives the number of Fock states of  $N$  bosonic particles distributed over, respectively,  $\mathcal{K}$  and  $\mathcal{K} - 1$  modes.

## VII Equivalent description of a lossy linear network

Below we focus on bosonic particles, for fermions one should replace the commutators below by the anti-commutators. A realistic network  $U$  is non-unitary due to (generally path-dependent) losses of particles, its action is described not just by input and output mode operators,  $a_1, \dots, a_M$  and  $b_1, \dots, b_M$ , but also by some additional operators  $f_1, \dots, f_M$  accounting for losses:

$$a_k^\dagger = \sum_{l=1}^M U_{k,l} b_l^\dagger + f_k^\dagger, \quad (89)$$

where operators  $f_k$  and  $f_k^\dagger$  commute with the creation and annihilation operators corresponding to network modes:  $[f_k, a_l] = [f_k, b_l] = [f_k^\dagger, a_l] = [f_k^\dagger, b_l] = 0$  [21]. Using the latter we obtain

$$[f_k, f_j] = 0, \quad [f_k, f_j^\dagger] = \delta_{k,j} - \sum_{l=1}^M U_{k,l}^* U_{j,l}. \quad (90)$$

Eq. (90) can be satisfied if we expand the loss operators  $f_1^\dagger, \dots, f_M^\dagger$  over some additional creation operators

$$f_k^\dagger = \sum_{l=1}^M V_{k,l} b_{M+l}^\dagger, \quad (91)$$

where  $V$  is a matrix satisfying the following matrix equation (valid both for bosons and fermions)

$$VV^\dagger = I - UU^\dagger, \quad (92)$$

where  $I$  is the unit matrix and  $U^\dagger$  denotes the Hermitian conjugate to matrix  $U$ . Notice that Eq. (92) requires that the singular values of  $U$  be bounded by 1, which is the necessary and sufficient condition for an arbitrary complex matrix  $U$  to describe a passive linear quantum network. There is a polar decomposition,  $U = \sqrt{A}\mathcal{U}$ , where  $\mathcal{U}$  is a unitary matrix and  $A = UU^\dagger$  a Hermitian matrix (describing losses in the network) with the eigenvalues bounded by 1.

The expansion in Eq. (91) means that one can imbed an arbitrary (non-unitary) linear  $M$ -mode network into a  $2M$ -mode unitary one [27]. The following embedding unitary network seems to be the simplest one

$$\hat{U} = \begin{pmatrix} U & V \\ -V^\dagger \mathcal{U} & D \end{pmatrix}, \quad (93)$$

here the diagonal matrix  $D = \text{diag}(\eta_1, \dots, \eta_M)$ ,  $0 \leq \eta_k \leq 1$ , is composed of the square-roots of singular values of  $U$  (eigenvalues of  $\sqrt{A}$ , since  $A = UU^\dagger$ ),  $V = SQ$ , with the unitary matrix  $S$  containing the eigenvectors of  $\sqrt{A}$ ,  $\sqrt{A} = SDS^\dagger$ ,  $Q = \text{diag}(\sqrt{1 - \eta_1^2}, \dots, \sqrt{1 - \eta_M^2})$ , and  $\mathcal{U}$  is from the polar decomposition  $U = \sqrt{A}\mathcal{U}$ . Matrix  $\hat{U}$  can be also rewritten in a product form

$$\hat{U} = \begin{pmatrix} S & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} D & Q \\ -Q & D \end{pmatrix} \begin{pmatrix} S^\dagger & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \mathcal{U} & 0 \\ 0 & I \end{pmatrix}, \quad (94)$$

from which it is evident that  $\hat{U}$  is unitary. Note that in description of a lossy network  $U$ , the unitary matrix  $\hat{U}$  has vacuum input in the modes  $\{M + 1, \dots, 2M\}$  and output modes  $\{M + 1, \dots, 2M\}$  are not accessible “loss channels”. The embedding matrix of Eq. (93) reduces to a matrix appeared in Ref. [22] in the special case of path-independent losses, i.e., a diagonal loss matrix  $A$ .

- 
- [1] C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert, arXiv:1306.3995 [quant-ph].
  - [2] S. Aaronson and A. Arkhipov, arXiv:1309.7460 [quant-ph].
  - [3] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Molmer, Phys. Rev. Lett. **113**, 020502 (2014).

- [4] M. C. Tichy, M. Tiersch, F. Mintert, and A. Buchleitner, New Journal of Phys. **14**, 093015 (2012).
- [5] V. S. Shchesnovich, Phys. Rev. A **91**, 013844 (2015).
- [6] S. Rahimi-Keshari *et al*, Optics Express **21**, 13450 (2013).
- [7] A. Leverrier and R. García-Patrón, Quant. Inform. Comput. **15**, 489 (2015).

- [8] A. Arkhipov, Phys. Rev. A **92**, 062326 (2015).
- [9] V. S. Shchesnovich, Phys. Rev. A **89**, 022333 (2014).
- [10] V. S. Shchesnovich, Int. J. Quantum Inform. **13**, 1550013 (2015).
- [11] E. Fisher, Archiv d. Math. u. Phys. **13**, 32 (1907).
- [12] E. H. Lieb, J. Math. and Mech. **16**, 127 (1966).
- [13] R. Merris, *Multilinear Algebra* p. 216 (Gordon and Breach Pubs. 1997).
- [14] D. Weingarten, J. Math. Phys. **19**, 999 (1978).
- [15] S. Aaronson and A. Arkhipov, Theory of Computing **9**, 143 (2013).
- [16] P. W. Brouwer and C. W. J. Beenakker, J. Math. Phys. **37**, 4904 (1996).
- [17] R. P. Stanley, *Enumerative Combinatorics*, 2nd ed., Vol. 1 (Cambridge University Press, 2011).
- [18] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integraly i Ryady. Elementarnue funktsii* (in Russian) (Moscow, FizMatLit 1981).
- [19] H. Minc, *Permanents, Encyclopedia of Mathematics and Its Applications*, Vol. **6** (Addison-Wesley Publ. Co., Reading, Mass., 1978).
- [20] S. Aaronson and A. Arkhipov, Theory of Computing **9**, 143 (2013).
- [21] S. M. Barnett, C. R. Gilson, B. Huttner, and N. Imoto, Phys. Rev. Lett. **77**, 1739 (1996).
- [22] S. Rahimi-Keshari, M. A. Broome, R. Fickler, A. Fedrizzi, T. C. Ralph, and A. G. White, Opt. Express **21**, 13450 (2013).
- [23] The authors of Ref. [3] do not claim that their assessment protocol is understood as a holistic one, it is used only to illustrate appearance of serious loopholes if one considers the protocol as holistic.
- [24] To rule out such and other possible networks is precisely the point in an *independent* certification of a network.
- [25] Since only polynomial number of runs of a BS device can be employed in an efficient assessment protocol, only a polynomial in  $N^{-1}$  accuracy in probability  $p_N^{(B)}$  can be achieved in an experiment.
- [26] Generally, given a p.s.d. Hermitian matrix  $A$ , representable in the form  $A = U\Lambda U^\dagger$  with  $U$  being a  $M$ -dimensional unitary matrix and  $\Lambda = \text{diag}(1, \dots, 1, 0, \dots, 0)$ , with  $\mathcal{K}$  ones, where  $M = O(N^{2+\delta})$  for some  $\delta > 0$  and  $M - \mathcal{K} = O(N)$ , and a preset probability of success scaling polynomially in  $N^{-1}$ , the described algorithm returns  $(1 + O(\epsilon))\text{per}(H)$  with the preset probability for any  $N$ -dimensional principal submatrix  $H$  of  $A$  in a number of flops scaling polynomially in both  $N$  and  $1/\epsilon$ .
- [27] Moreover, analysis of necessary conditions on such embedding shows that one cannot reduce the size of an embedding network for non-singular matrices  $U$ .